

Proposal for Characterization of 3DTV Video Sequences Describing Parallax Information

J. P. López, J. A. Rodrigo, D. Jiménez, J. M. Menéndez

Abstract—Recommendations such as P.910 suggests parameters TI (temporal information) and SI (spatial information) for characterizing video sequences for quality assessment. In this paper, we suggest two additional parameter based on disparity called SPI (spatial parallax information) and TPI (temporal parallax information) to characterize 3DTV video sequences for this purpose.

I. INTRODUCTION

Characterization of video sequences is a key factor when developing video quality assessment. Recommendation P. 910 [1] widens the horizons from previous recommendation BT.500 [2] including two parameters for defining video sequences, which are necessary to demonstrate the validity of the material for the tests: Spatial perceptual information (SI) and temporal perceptual information (TI).

These two parameters are necessary to justify the usage of every individual video sequence included in an objective or a subjective evaluation test, but they are insufficient when the environment is 3DTV stereoscopic video. 2D methods have severe limitations when analyzing stereoscopy [3]. Recommendation BT.2021 [4] offers specific requirements for stereoscopic quality tests. These requirements are specifically focused on analyzing the observer's correct vision, when involved in subjective assessment, but it lacks of information about characterization of video sequences used in the studies.

VQEG is currently working on defining new models for 3D Video Quality Assessment [5]. Other proposals of new requirements for subjective assessment are discussed in [6], including depth rendering abilities and limitations of conventional 2D recommendations, after understanding new characteristics by 3DTV video sequences. Depth rendering and analysis of different multiview displays are part of the new considerations to have in this special type of video assesment, but video characterization is reduced to spatio-temporal parameters defined in previous standards.

Low complexity algorithms such as Sum of Absolute Differences (SAD) can perform well calculating parallax diagrams and depth maps under most circumstances, as stated by Scharstein et al [7]. Parallax is necessary to define the stereoscopic complexity of each video sequence and it results useful for quality assessment [8]. From parallax estimation, both parameters Spatial Parallax Information (SPI) and Temporal Parallax Information (TPI) are obtained, offering extra information regarding the sequence characterization.

By calculating SPI and TPI parameters, new three-dimensional plots can be represented. By adding these two new parameters to the classic TI vs. SI representation, a more complete characterization of the sequences is achieved. Thus,

the analysis of the validity of video sequences databases is improved.

II. PARAMETERS SPI AND TPI

When selecting test sequences, it can be useful to compare the relative spatial and temporal information with the stereoscopic component found in the set of sequences available. In this way, the selection of the sequences may be more reliable and, accordingly, the number of sequences may be reduced which is essential when defining subjective assessment tests.

Selected sequences should span the highest portion of the spatio-temporal-depth information plane, covering the majority of cases with different combinations of motion and spatial detail, and also with parallax variations.

A. Spatial Parallax Information (SPI)

The spatial parallax information is based on the depth variation in every single frame of the sequence. Parallax diagram for each video frame (luminance plane) at time n (F_n). SAD algorithm is used to calculate parallax diagrams to evaluate depth variation in the sequence as explained in Section IV, obtaining the parallax function $P[F_n(i, j)]$ over luminance component for every single pixel.

The standard deviation over space (std_{space}) in parallax diagram of each frame is then computed, after all frames in the sequence are computed. The maximum value for a frame in the complete time sequence is defined to represent the parallax spatial information content of the scene. In equation form, the SPI parameters can be defined as follows.

$$SPI = \max_{time} \{std_{space}[P(F_n)]\}$$

B. Temporal Parallax Information (TPI)

The temporal parallax information parameter is based upon the depth difference function $D_n(i, j)$, which is the difference between the values of parallax of the pixels $P[F_n(i, j)]$ when calculating SAD function at the same location in space at successive frames. Being $F_n(i, j)$ the value of luminance in each pixel from frame n , depth function is defined as:

$$D_n(i, j) = P[F_n(i, j)] - P[F_{n-1}(i, j)]$$

The measure of temporal parallax information (TPI) parameter is computed as the maximum over time (\max_{time})

of the standard deviation over space std_{space} of $D_n(i, j)$ over each i and j

$$TPI = \max_{time} \{std_{space}[D_n(i, j)]\}$$

More parallax variation in adjacent frames will result in higher values of TPI.

III. GRAPHICAL REPRESENTATION

The graphical representation is a bubble plot in which the bubbles are ellipses whose axes define the parameters related to parallax information.

In Rec. P.910 [1] the SI parameter is defined as a measure that defines the amount of spatial detail of a sequence, being higher when scenes are more spatially complex. On the other hand, TI parameter is defined as a measure that indicates temporal changes of the video sequence, defining high motion sequences. In this recommendation examples of spatio-temporal plots are collected to see the distribution of sequences derived from these two parameters.

As seen in Fig. 1, the position of the bubble in spatial coordinates defines the traditional values of TI and SI, with the spatial and temporal information of the video sequence. For 3D stereoscopic contents, each video sequence is more accurately defined with two additional parameters which correspond to spatio-temporal variations of parallax. These two parameters have been described in the previous section.

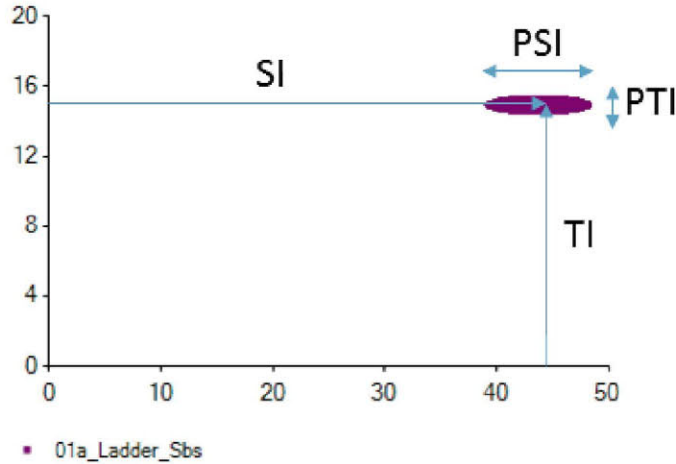


Fig. 1. Example of the four parameters defined for each sequence in the diagram

IV. PARALLAX DIAGRAM

The system performs a SAD based algorithm for computing depth maps and parallax diagrams from each stereoscopic pair of images (both views) extracted from a side-by-side video sequence. SAD algorithm is chosen because it is one of the simplest algorithms but offers satisfactory results. SAD based algorithms are less computational demanding and are able to fulfill the goal of providing the 3D complementary parameters.

The most difficult detection with SAD algorithm occurs in homogeneous zones, where the capability to discern between possible pair candidates is low. For this purpose, the system performs a difference between both views, only over those pixels that differs from one view to the other, avoiding noise in the resulting depth maps. Discarded pixels won't be taken into account for statistical calculations. The histogram will contain less depth information, although is more accurate. Fig.2 shows the original depth map (left) with SAD algorithm and the filtered depth map (right). In the original depth map there are several errors in the background zone, where the sky is homogeneous. Filtering in depth map improves the accuracy for the definition of parallax diagram.

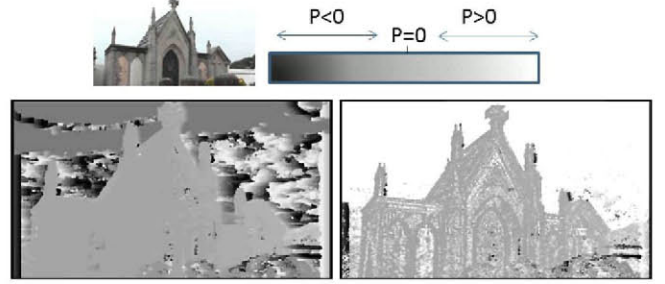


Fig.2. Depth map. (P refers to parallax)

Fig. 3 shows the histogram calculated for Fig.2. All the elements in the scene have positive parallax. There is a very small amount of negative parallax pixels which represent noise (bad information) that result from the depth map algorithm calculation.

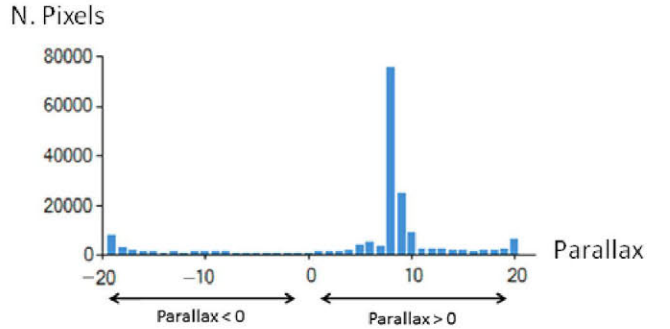


Fig. 3. Example of parallax histogram corresponding

V. RESULTS AND EXAMPLES

Fig. 4 shows an example of the plot which includes the spatial, temporal and parallax information of a collection of stereoscopic video sequences [9]. The TI axis contains the temporal information, with the still scenes located around 0, and near the top of this axis sequences with a lot of motion are found. Along the SI axis, the spatial information is shown. It varies from homogeneous pictures that are closer to the intersection to high-detailed sequences located near the top of this axis.

The third and fourth components which are included in the plot correspond to the axes of the elliptical bubble. The higher the value in the axis is, the higher the variation of spatial and temporal parallax values in that sequence. Then, analyzing

Fig. 4, we can conclude that sequences as “Itaca3D” and “Palco3D” offer the highest values of TPI, and also a high SPI. On the other hand, sequence “Muscles” offer the lowest SPI and TPI from this set.

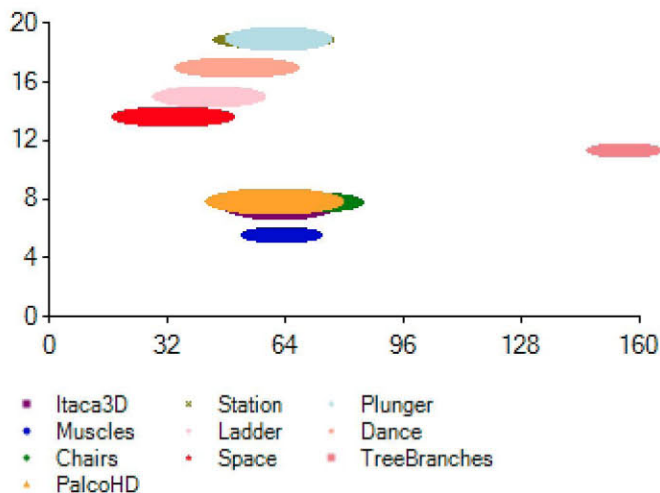


Fig. 4. Spatio-temporal and spatial-parallax plot for ten 3DTV sequences

Most of the sequences included in the diagram come from NAMA3DS1-COSPAD1 database [9].

VI. CONCLUSIONS

In this paper, two parameters for stereoscopic 3DTV video sequences has been proposed and discussed. New requirements for parallax description in the spatial and temporal fields are necessary for evaluating content disparity affecting the depth representation. The description of parameters SPI and TPI have been explained with illustrated graphs in the paper.

In order to produce a reliable and comparable subjective experiment, test material should be carefully selected and result analysis should cover the higher range of characteristics as possible. Compared to 2D assessment, 3DTV shows its diversity and complexity. For this purpose, the inclusion of these parameters recommendation in an international standardization organization is highly recommended, in order to widen video characterization for 3DTV video quality assessment. Additionally, the exportation of results derived from them to different environments would be facilitated. Our future work is focused on this evaluation and experiments over this characterization as the basis of quality metrics.

ACKNOWLEDGMENT

The work developed in Universidad Politécnica de Madrid performed in the framework of project TEC2012-38402-C04-01 HORFI, which is partially funded by the Spanish Ministry of Science and Innovation. The authors would like to acknowledge the Spanish broadcaster RTVE, the manufacturer company Philips for lending the television used in subjective assessment and University of Nantes for the cession of the stereoscopic database.

REFERENCES

- [1] Recommendation ITU-T P.910 (v.04/2008). Telecom. Standardization Sector OF ITU “Subjective video quality assessment methods for multimedia applications”.
- [2] Recommendation ITU-R. (2002). BT. 500-11, Methodology for the subjective assessment of the quality of television pictures, 22, 25-34.
- [3] E. Dumić, S. Grgić, D. Jiménez and L.A. da Silva Cruz, Benchmark of State of the Art Objective Measures for 3D Stereoscopic Video Quality Assessment on the Nantes Database. ELMAR 2014
- [4] Recommendation ITU-R, Rec. BT. 1438. Subjective assessment of stereoscopic television pictures, 2000. Recommendation ITU-R BT. 2021. “Subjective Assessment of Stereoscopic 3DTV Systems (08/2012)”, 2012.
- [5] “Establishing a Ground Truth for Quality of Experience in 3D for assessment methodologies in 3D Video Quality Assessment”. Available at: <http://www.its.bldrdoc.gov/vqeg/projects/3dtv/3dtv.aspx>
- [6] Chen, W., Fournier, J., Barkowsky, M., & Le Callet, P. (2010). New requirements of subjective video quality assessment methodologies for 3DTV. 2010.
- [7] D. Scharstein, R. Szeliski. “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms”, International Journal of Computer Vision, No 47, June 2002.
- [8] López, Juan Pedro; Rodrigo, Juan Antonio; Jiménez, David; Menéndez, José Manuel; „Stereoscopic 3D video quality assessment based on depth maps and video motion, EURASIP Journal on Image and Video Processing, 2013, 1, 1-14, 2013, Springer International Publishing
- [9] Urvoy, M., Barkowsky, M., Cousseau, R., Koudota, Y., Ricorde, V., Le Callet, P., et al. (2012, July). NAMA3DS1-COSPAD1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3D stereoscopic sequences. In Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on (pp. 109-114). IEEE.